Least Costly Identification Experiment for the Identification of One Module in a Dynamic Network

Federico Morelli, Xavier Bombois, Håkan Hjalmarsson, Laurent Bako, Kévin Colin

Least Costly Identification Experiment for the Identification of One Module in a Dynamic Network

Federico Morelli [a] Xavier Bombois [a,b] Håkan Hjalmarsson [c],
Laurent Bako [a] Kévin Colin [c]

[a] *Laboratoire Ampère, Ecole Centrale de Lyon, Université de Lyon, Ecully, France*

[b] *Centre National de la Recherche Scientifique (CNRS), France*

[c] *Automatic Control, School of Electrical Engineering, KTH, Stockholm, Sweden*

Abstract

In this paper we consider the design of least costly experiments for the identification of one module in a given network of locally controlled systems. The identification experiment will be designed in such a way that we obtain a sufficiently accurate model of the to-be-identified module with the smallest identification cost i.e. with the least perturbation of the network.

Key words: Optimal identification Experiment Design; Dynamic Network Identification

1 Introduction

This paper contributes to the efforts of developing techniques for the identification of large-scale or interconnected systems when the topology of the network is known. In many papers, the problem is seen as a multivariable identification problem and structural properties of the system are then used to simplify this complex problem (see e.g. [7]). The identifiability of the multivariable structure is studied in a prediction error context in [18] while this multivariable structure is exploited in other papers to reduce the variance of a given module in the network (see [8,4]). In other contributions, conditions are derived for the consistent estimation of a given module in a dynamic network (see e.g. [3,6]).

While many different problems have thus been extensively studied in the dynamical network context, this is not the case for optimal experiment design (i.e. the problem of designing the excitation signal of an identification experiment to guarantee a certain model accuracy under some constraints on this excitation signal). In our previous contribution [1], we made the first steps towards optimal experiment design in a dynamical network context. In [1], we considered the case of a network made up of locally controlled systems, i.e. modules, whose interconnection is realized by exchanging their measured output between neighbouring modules (this type of networks is usual in the literature on multi-agent systems (see e.g. [5,11])). For this particular type of dynamical networks, we showed how to design the excitation signals that have to be added to each module in order to identify models of these different modules that are sufficiently accurate to enhance the network performance by a redesign of the local controllers. The accuracy of each model can be measured by the inverse of the covariance matrix of the identified parameter vector of each module. In [1], we have derived an expression for the inverse of this covariance matrix as an affine function of the excitation signal spectra.

Note that [10] illustrates the identification methodology introduced in [1] via a realistic simulation example (i.e. a platoon of autonomous cars).

In this paper, like in [1], we will consider the optimal experiment design problem for a network of locally controlled systems. However, unlike in [1], we will do that for the case where we are only interested in the accurate identification of one specific module l of the network. Like in [1], to maintain the network performance, this identification will be performed in the original network configuration via the application of an excitation signal to module l. The contribution of the present paper is the extension of the least costly identification framework (see [2]) to this particular dynamic network identification problem. In particular, we design the spectrum of the excitation signal applied to l in such a way that the accuracy of the identified model (measured via the inverse of the covariance matrix) is larger than a given threshold while entailing the smallest perturbation on

the network. The perturbation (i.e. the cost of the identification) will be measured by the sum of the effects of the excitation signal on the input and output of each system in the network.

With respect to the least costly framework introduced in [2] for a single closed loop, the cost of the identification experiment in the network context thus not only contains the perturbation induced by the excitation signal in the closed loop where the system has to be identified, but also the perturbation induced in other loops by this excitation signal. This propagation of the effect of the excitation signal is due to the fact that the output signal of the to-be-identified loop (which is perturbed by the excitation signal) is transmitted to neighbouring modules. In this paper, in order to reduce this propagation, we propose an approach where the signal transmitted to the neighbouring modules is no longer the actual output signal, but a sanitized version of this output signal where the contribution of the excitation signal has been (partially) removed. Indeed, using an initial estimate of the to-be-identified system, we derive an estimate of this contribution and we subtract this estimate from the measured output signal before the transmission to the neighbouring modules.

This new configuration is inspired by the concept of stealth identification that we introduced in [15] for a single closed loop and that we here extend to the network case. The use of the stealth identification in this paper is also a new application of this concept since, in [15], it was introduced as a tool to enable classical optimal experiment design in a loop where the controller is not Linear Time Invariant (LTI). With respect to [15], we also analyze which accuracy condition the initial estimate used to compute the sanitized version of the output signal must respect for the stealth configuration to be effective (i.e., to yield a smaller identification cost).

Notations: The matrix I_n denotes the identity matrix of dimension n. The matrix of dimension $N \times N$

$$\begin{pmatrix} X_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & X_N \end{pmatrix}$$

will be denoted $diag(X_1, ..., X_N)$. For a matrix A, A^T denotes the transpose of A and A^* its conjugate transpose. For a vector of transfer functions $R(z)$, $R_i(z)$ denotes the i^{th} entry of $R(z)$. In addition, the symbol z will not only represent the Z-transform variable, but also the shift operator. Finally, $\otimes$ denotes the Kronecker product and E the expectation operator.

2 Description of the network configuration

We consider a network made up of N_{mod} single-input single-output (SISO) systems $\mathcal{S}_i$ $(i = 1, \ldots, N_{mod})$ operated in closed loop with a SISO decentralized controller K_i $(i = 1, \ldots, N_{mod})$:

$$\mathcal{S}_i : \; y_i(t) = G_i(z, \theta_{0,i})u_i(t) + v_i(t) \quad (1)$$
$$u_i(t) = K_i(z)(y_{ref,i}(t) - y_i(t)) \quad (2)$$

where the signal u_i is the input applied to the system $\mathcal{S}_i$ and y_i is the measured output. This output is made up of a contribution of the input u_i and of a disturbance term $v_i(t) = H_i(z, \theta_{0,i})e_i(t)$ that represents both process and measurement noises. The different systems $\mathcal{S}_i$ $(i = 1, \ldots, N_{mod})$ are thus described by two stable transfer functions $G_{0,i}(z) = G_i(z, \theta_{0,i})$ and $H_{0,i}(z) = H_i(z, \theta_{0,i})$, the latter being also minimum-phase and monic. These transfer functions are parametrized by an unknown true parameter vector $\theta_{0,i} \in \mathbf{R}^{n_{\theta_i}}$. For each i $(i = 1, \ldots, N_{mod})$, the signal e_i $(i = 1, \ldots, N_{mod})$ defining v_i is a zero mean white noise signal of variance $\sigma^2_{e_i}$. The variances $\sigma^2_{e_i}$ $(i = 1, \ldots, N_{mod})$ are thus the diagonal elements of the covariance matrix Σ of the vector $\bar{e} = (e_1, e_2, ..., e_{N_{mod}})^T$ that is assumed strictly positive definite (i.e., $E\bar{e}(t)\bar{e}(t)^T = \Sigma > 0$). We further assume that $E\bar{e}(t)\bar{e}(t-\tau)^T = 0$ for all $\tau \neq 0$. Finally, in (2), $y_{ref,i}$ is a reference signal that will be computed based on the measured outputs of neighbouring modules (see later). We can rewrite the above equations as follows:

$$\bar{y}(t) = \bar{G}(z, \theta_0)\bar{u}(t) + \bar{H}(z, \theta_0)\bar{e}(t) \quad (3)$$
$$\bar{u}(t) = \bar{K}(z)(\bar{y}_{ref}(t) - \bar{y}(t)) \quad (4)$$

where $\bar{y}$, $\bar{u}$, $\bar{y}_{ref}$ are defined in a similar way as $\bar{e}$ and where $\theta_0 = (\theta^T_{0,1}, \ldots, \theta^T_{0,N_{mod}})^T \in \mathbf{R}^{n_\theta}$ concatenates the true parameter vectors $\theta_{0,i}$ $(i = 1, ..., N_{mod})$. In these equations, we also use the notation $\bar{G} = diag(G_1, \ldots, G_{N_{mod}})$ ($\bar{H}$ and $\bar{K}$ are defined in a similar way).

The closed-loop systems described in (3)-(4) are interconnected via the following equation:

$$\bar{y}_{ref}(t) = \mathcal{A}\, \bar{y}(t) \; + \; \mathcal{B}\, ref_{ext}(t) \quad (5)$$

where the matrix $\mathcal{A}$ and the vector $\mathcal{B}$ represent the flow of information in the network and ref_{ext} is a (scalar) external reference signal that should be followed by all outputs y_i and that is generally only available at one node of the network. This type of interconnections is typical in formation control or multi-agent systems (see e.g. [5,11]).

To illustrate (5), let us consider the network represented in Figure 1. In this network, we have $N_{mod} = 6$ systems/modules, all of the form (1) and all operated as in (2) with a decentralized controller K_i (see Figure 2). These local closed loops are represented by a circle/node in Figure 1. The objective of this network is that the outputs y_i of all modules follow the external reference ref_{ext} even though this reference is only available at Node 1. For this purpose, a number of nodes are allowed to exchange information (i.e. their measured output) with

some other neighbouring nodes. The arrows between the nodes in Figure 1 indicate the flow of information.

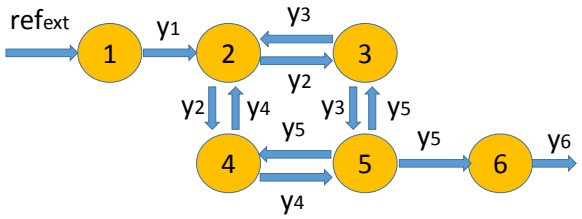

Fig. 1. Example of graph representation of the network, each circle represents a node i and the edges represent the communication link between the nodes

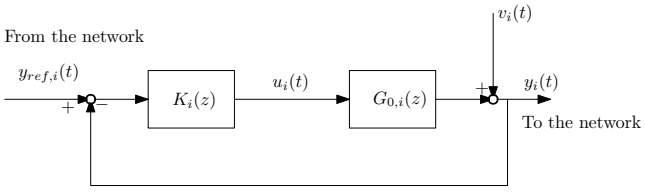

Fig. 2. Representation of a single module/node i

For example, Node 5 receives the output of two nodes (i.e. Nodes 3 and 4) and sends its output (i.e. y_5) to three nodes (Nodes 3, 4, and 6). The reference signal $y_{ref,i}$ of Node i will be computed as a linear combination of the received information at Node i. For Node 5, $y_{ref,5}$ will thus be a linear combination of y_3 and y_4. More precisely, for all outputs y_i to be able to follow the external reference ref_{ext}, $\mathcal{A}$ and $\mathcal{B}$ in (5) are generally chosen as [5,11]:

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathcal{B} = (1, 0, ..., 0)^T.$$

The matrix $\mathcal{A}$ is called the normalized adjacency matrix in the literature [5]. Using (5), we, e.g., see that the tracking error signals $y_{ref,1} - y_1$ and $y_{ref,2} - y_2$ of Nodes 1 and 2 are respectively given by $ref_{ext} - y_1$ and $1/3\left((y_1 - y_2) + (y_3 - y_2) + (y_4 - y_2)\right)$. Similar relations can be found for all the other nodes. If the different loops $[K_i\ G_i]$ are designed to make the tracking error $y_{ref,i} - y_i$ as small as possible, it can be proven that such an interconnection allows good tracking of ref_{ext} at all nodes [11,5]. A normalized adjacency matrix can be defined for any information flow using the following rules. Row i of $\mathcal{A}$ is zero if no output is sent to Node i. If y_i is sent to Node j, $\mathcal{A}_{ji}$ (i.e., the entry (j,i) of $\mathcal{A}$) will be nonzero. Finally, all nonzero entries in a row of $(\mathcal{A}, \mathcal{B})$ are equal and sum up to one.

We also need to introduce the notion of *(directed) path* between two nodes. There exists a path from Node i to Node j if $\mathcal{A}_{ji} \neq 0$ or we can find a set of ζ intermediary nodes described by the indexes $\{n_1, ..., n_\zeta\}$ such that $\mathcal{A}_{n_1 i} \neq 0$, $\mathcal{A}_{n_2 n_1} \neq 0$, ... , $\mathcal{A}_{j n_\zeta} \neq 0$. Using this notion of path, Definition 1 introduces a set of indexes for each node of the network:

Definition 1 *Consider an arbitrary node of a network containing N_{mod} nodes, say Node j ($j = 1, ..., N_{mod}$). For this node, we define the set $\mathcal{P}_j$ as the set of indexes $i \neq j$ such that there is a path from Node j to Node i.*

As an example, $\mathcal{P}_5 = \{2, 3, 4, 6\}$ for the network of Figure 1. For the sequel, it is important to note the following fact. If an external signal (e.g., an excitation signal r) is added to Node j, this external signal will also influence all nodes i with $i \in \mathcal{P}_j$.

Remark. Among the networks of the type described above, a special, but frequent situation is the *homogenous network* situation where the systems $\mathcal{S}_i$ in (1) are all identical i.e., the parameter vectors $\theta_{0,i}$ ($i = 1, ..., N_{mod}$) are all equal. The network of autonomous cars considered in [10] or the network of Phase-Locked Loops (PLL's) considered in [11] are examples of such homogenous networks.

3 Identification of one module in the network and cost of the experiment

3.1 Objective

In this paper, we wish to obtain, via prediction error identification, an accurate estimate of a single node in the network, say Node l. In other words, we wish to obtain an accurate estimate of the parameter vector $\theta_{0,l}$ describing Node l (we will precisely define the desired level of accuracy in Section 3.3). We may be interested by an accurate model of this single node for different reasons. Node l can for example be critical for the performance of the network and it is therefore important to monitor its dynamics via frequent re-identification to be able to update the control strategy if this dynamics has changed. Another case where this problem can occur is the case where we wish to improve the control performance of an homogenous network. In this case, identifying one node with sufficient accuracy is equivalent to identifying all the nodes with sufficient accuracy and therefore allows to redesign the local controllers to improve the control performance of the network.

We can obtain an accurate estimate of $\theta_{0,l}$ in different ways. A possible way would be to disconnect Node l from the network and perform the identification using classical open-loop or closed-loop identification. However, Node l may be essential for the network performance and taking such an action may strongly affect this performance and, if not done with care, even destabilize the network. Disconnecting Node l indeed means changing

$\mathcal{A}$ so that the interconnection (3)-(4)-(5) may become unstable. Therefore, in this contribution, we will keep the interconnection structure (3)-(4)-(5) intact. By this, we mean that the transfer functions between ref_{ext} and the node signals y_i and u_i $(i = 1, ..., N_{mod})$ will not be modified, therefore maintaining the tracking ability of the network and we also mean that we will not modify the transfer functions between the disturbances v_i and the node signals y_i and u_i $(i = 1, ..., N_{mod})$, thus leaving the disturbance rejection performance of the network intact. In order to obtain informative data for the identification of an accurate estimate of $\theta_{0,l}$ while, at the same time, keeping these transfer functions unaltered, we can inject external/exogenous excitation signals at well-chosen locations (similarly as what is done in closed-loop identification [13]). Consequently, with respect to the normal operation (3)-(4)-(5), the only perturbation during the identification experiment will be the perturbation induced by these external excitation signals and this perturbation (the cost of the experiment) will be minimized by an appropriate design of these excitation signals.

In the sequel, we will suppose that an initial estimate $\theta_{init,l}$ of $\theta_{0,l}$ is available, but that this initial estimate is not sufficiently accurate (otherwise, the identification described above would be useless). This initial estimate $\theta_{init,l}$ of $\theta_{0,l}$ defines initial estimates $G_l(z, \theta_{init,l})$ and $H_l(z, \theta_{init,l})$ of $G_l(z, \theta_{0,l})$ and $H_l(z, \theta_{0,l})$, respectively. More details on how this initial estimate is obtained will be given in Section 5.

3.2 Cost of an experiment in the stealth and non-stealth configurations

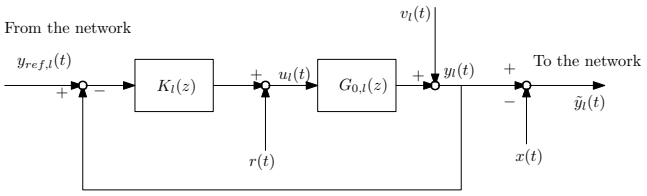

Fig. 3. To-be-identified node (i.e. Node l) during the identification experiment. In the stealth setting, $x(t)$ is given by (6). In the non-stealth setting, $x(t) = 0$.

During the identification experiment, we will apply for a duration N (i.e. from $t = 1$ till $t = N$) an external excitation signal $r(t)$ of spectrum Φ_r at the output of the controller K_l of Node l (see Figure 3). This external excitation (that we assume uncorrelated with ref_{ext} and $\bar{e}$) will allow to obtain sufficiently informative data $u_l(t)$ and $y_l(t)$ $(t = 1, \ldots, N)$ for an accurate identification of $\mathcal{S}_l$ (see Section 3.3). In Figure 3, we also observe the signal $x(t)$. In this paper, we will consider two choices for $x(t)$ corresponding to two settings: the stealth and the non-stealth configurations. In the non-stealth setting, the signal $x(t)$ will be chosen equal to zero. This choice corresponds to the classical setting for an identification experiment in a closed-loop/network context (see e.g. [1]). In the stealth setting, the signal $x(t)$ is chosen as the following estimate of the contribution of r to y_l:

$$x(t) = \underbrace{\frac{G_l(z, \theta_{init,l})}{1 + K_l(z)G_l(z, \theta_{init,l})}}_{=T_{init,l(z)}} r(t) \quad (t = 1, \ldots, N) \quad (6)$$

where $\theta_{init,l}$ is the initial estimate of $\theta_{0,l}$ (see Section 3.1). Note that (6) can be easily computed since both r and $T_{init,l}(z)$ are known.

As shown in Figure 3, the signal $x(t)$ is subtracted from the measured output $y_l(t)$ to give $\tilde{y}_l(t) = y_l(t) - x(t)$ which will be the signal that will be transmitted to compute $\bar{y}_{ref}$. Consequently, during the identification experiment, the equations (4)-(5) become:

$$\bar{u}(t) = \bar{m}_l \; r(t) + \bar{K}(z) \; (\bar{y}_{ref}(t) - \bar{y}(t)) \quad (7)$$
$$\bar{y}_{ref}(t) = \mathcal{A} \; (\bar{y}(t) - \bar{m}_l x(t)) + \mathcal{B} \; ref_{ext}(t) \quad (8)$$

where $\bar{m}_i$ $(i = 1, \ldots, N_{mod})$ denotes a unit (column) vector of dimension N_{mod} for which the i^{th} entry is equal to 1 and the other entries are equal to zero.

We will show, in the sequel, the advantage of the stealth setting in order to reduce the cost of the identification experiment. In the stealth setting, the output vector $\bar{y}$ and the input vector $\bar{u}$ in the network configuration (3)-(7)-(8) can be rewritten as follows as a function of the external signals r, ref_{ext} and $\bar{e}$:

$$\bar{y}(t) = R^y(z, \theta_0)r(t) + R^y_{ext}(z, \theta_0)ref_{ext}(t) + S^y(z, \theta_0)\bar{e}(t) \quad (9)$$
$$\bar{u}(t) = R^u(z, \theta_0)r(t) + R^u_{ext}(z, \theta_0)ref_{ext}(t) + S^u(z, \theta_0)\bar{e}(t) \quad (10)$$

for some vectors of transfer functions R^y_{ext}, R^u_{ext}, some matrices of transfer functions S^y, S^u and

$$R^u(z, \theta_0) = \bar{m}_l \; S_{0,l}(z) + \mathcal{N}(z, \theta_0) \; (T_{0,l}(z) - T_{init,l}(z)) \quad (11)$$
$$R^y(z, \theta_0) = \bar{G}(z, \theta_0) \; R^u(z, \theta_0) \quad (12)$$

where $S_{0,l}(z) = 1/(1 + K_l(z)G_l(z, \theta_{0,l}))$ and $T_{0,l}(z) = G_l(z, \theta_{0,l})/(1 + K_l(z)G_l(z, \theta_{0,l}))$ are scalar transfer functions and $\mathcal{N}(z, \theta_0)$ is a vector of transfer functions:

$$\mathcal{N}(z, \theta_0) = \bar{K}(z)\bar{S}(z, \theta_0)\Lambda(z, \theta_0)\mathcal{A} \; \bar{m}_l \quad (13)$$

with $\bar{S}(z, \theta_0) = \left(I_{N_{mod}} + \bar{G}(z, \theta_0)\bar{K}(z)\right)^{-1}$ and $\Lambda(z, \theta_0) = \left(I_{N_{mod}} - \mathcal{A}\bar{S}(z, \theta_0)\bar{G}(z, \theta_0)\bar{K}(z)\right)^{-1}$. For the sequel, it is important to note that $T_{init,l}(z)$ (see (6)) is the initial model of $T_{0,l}(z)$ that corresponds to the initial estimate $\theta_{init,l}$. Consequently, in (11), $T_{0,l}(z) - T_{init,l}(z)$ is the modeling error of this initial model $T_{init,l}(z)$.

In the normal operation described by (3)-(4)-(5), the vector $\bar{y}$ (resp. $\bar{u}$) is given by (9) (resp. (10)) with $r = 0$. Consequently, the tracking property of the network (measured by the transfer vectors R^y_{ext} and R^u_{ext}) and its

disturbance rejection property (measured by the transfer matrices S^y and S^u) are maintained during the identification experiment. The proposed experiment thus respects the objectives presented in Section 3.1 and we see that, with respect to the normal operation, the only perturbation of the vectors $\bar{y}$ and $\bar{u}$ during the identification experiment is $R^y(z)r(t)$ and $R^u(z)r(t)$, respectively. Consequently, it makes sense to define the cost of the identification experiment as the following function of the spectrum Φ_r of the excitation signal r:

$$J(\Phi_r, \theta_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{R}_\eta^*(e^{j\omega}, \theta_0) \mathcal{R}_\eta(e^{j\omega}, \theta_0)\ \Phi_r(\omega)\ d\omega \tag{14}$$

where $\mathcal{R}_\eta(z, \theta_0)$ is a vector of transfer functions given by $\mathcal{R}_\eta(z, \theta_0) = (R^y(z, \theta_0)^T,\ \sqrt{\eta}\ R^u(z, \theta_0)^T)^T$ with $\eta > 0$ a user-chosen weighting factor. The cost $J(\Phi_r, \theta_0)$ can be rewritten as $J(\Phi_r, \theta_0) = \sum_{i=1}^{N_{mod}} J_i(\Phi_r, \theta_0)$ where $J_i(\Phi_r, \theta_0)$ $(i = 1, ..., N_{mod})$ are the individual costs in each module:

$$\begin{aligned} &J_i(\Phi_r, \theta_0) = ... \\ &...\tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \left(|G_i(e^{j\omega}, \theta_{0,i})|^2 +\ \eta\right)\ |R_i^u(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega)\ d\omega \end{aligned} \tag{15}$$

with R_i^u being the i^{th} entry of R^u.

In the non-stealth setting, the cost $J(\Phi_r, \theta_0)$ of an identification experiment and the individual costs $J_i(\Phi_r, \theta_0)$ can be defined in a very similar way. However, the expressions for R^y and R^u that are used in (14) and (15) have a different expression:

$$R^{u,NS}(z, \theta_0) = \bar{m}_l\ S_{0,l}(z) + \mathcal{N}(z, \theta_0)\ T_{0,l}(z) \tag{16}$$

$$R^{y,NS}(z, \theta_0) = \bar{G}(z, \theta_0)\ R^{u,NS}(z, \theta_0) \tag{17}$$

By comparing (16) and (11), we observe that the modeling error $T_{0,l}(z) - T_{init,l}(z)$ is replaced by $T_{0,l}(z)$ in the expression of $R^{u,NS}$.

For both the stealth and non-stealth settings, the individual costs $J_i(\Phi_r, \theta_0)$ for nodes to which there is a path from Node l (i.e. those in $\mathcal{P}_l$) will be nonzero whereas, for nodes $i \neq l$ which are not in $\mathcal{P}_l$, the cost will be zero. Consequently, the excitation signal r will not have an influence only on Node l (where it is applied and where it is necessary for the identification of $\mathcal{S}_l$), but also on all nodes i with $i \in \mathcal{P}_l$. For the network in Figure 1, if the excitation signal r is applied in Node 5, besides $J_5(\Phi_r, \theta_0)$, the individual costs $J_i(\Phi_r, \theta_0)$ for $i = 2, 3, 4$, and 6 will also be non-zero. This result is equivalent to the fact that $\mathcal{N}_i(z, \theta_0)$ (i.e. the i^{th} entry of $\mathcal{N}(z, \theta_0)$) is a nonzero transfer function for all $i \in \mathcal{P}_l$ and is equal to zero for all $i \neq l$, $i \notin \mathcal{P}_l$.

The role of the stealth compensation $x(t)$ is to reduce this propagation of the excitation r (applied in Node l) to the nodes in $\mathcal{P}_l$. Before explaining this in more details, let us make the following assumptions on Node l and the considered network:

Assumption 1 *Consider the set $\mathcal{P}_l$ (see Definition 1) corresponding to the to-be-identified Node l. We assume that $\mathcal{P}_l$ is a non-empty set.*

Assumption 2 *Consider the set $\mathcal{P}_l$ (see Definition 1) corresponding to the to-be-identified Node l and the vector of transfer functions $\mathcal{N}(z, \theta_0)$ (see (13)). We assume that, for all $i \in \mathcal{P}_l$, the i^{th} entry $\mathcal{N}_i(z, \theta_0)$ of $\mathcal{N}(z, \theta_0)$ is such that $\mathcal{N}_i(e^{j\omega}, \theta_0) \neq 0$ for (almost) all frequencies.*

If $\mathcal{P}_l$ would be empty, there is of course no need for the stealth setting since the signal r will not be propagated to other nodes. Assumption 2 will in fact always hold, except in pathological cases that we here want to formally exclude.

We can now explain the role of the stealth compensation in reducing the propagation of the influence of the excitation r towards the nodes $i \in \mathcal{P}_l$. Let us first consider the ideal case i.e. when $T_{init,l} = T_{0,l}$. This choice does not change the situation in the non-stealth setting i.e. $J_i(\Phi_r, \theta_0)$ remains nonzero for all $i \in \mathcal{P}_l$ since $R_i^{u,NS}$ remains the same nonzero transfer function for all these i (it is not function of $T_{init,l}$). However, in the stealth setting, for all $i \neq l$, the transfer function R_i^u is identically zero when $T_{init,l} = T_{0,l}$. Consequently, in this ideal case, the effect of the excitation $r(t)$ will only be tangible in the to-be-identified module i.e. $J_i(\Phi_r, \theta_0) = 0$ for all $i \neq l$.

In practice, $T_{init,l}$ will of course always be different from $T_{0,l}$, but, as shown in the following proposition, the stealth configuration will remain beneficial under the mild condition that the relative error between $T_{init,l}$ and $T_{0,l}$ is less than 100%.

Proposition 1 *Consider that, following the procedure described in this section, an excitation signal $r(t)$ of spectrum Φ_r is applied to Node l of a network like the one described in Section 2 and satisfying Assumptions 1 and 2. Let us for this spectrum Φ_r compute the individual costs $J_i(\Phi_r, \theta_0)$ $(i = 1, ..., N_{mod})$ in the stealth setting and in the non-stealth setting using the respective expression for these costs in the two settings (see Section 3.2). Suppose finally that the initial model $T_{init,l}(z)$ of $T_{0,l}(z)$ satisfies the following accuracy constraint at the frequencies ω where $\Phi_r(\omega) \neq 0$:*

$$\frac{|T_{0,l}(e^{j\omega}) - T_{init,l}(e^{j\omega})|}{|T_{0,l}(e^{j\omega})|} < 1. \tag{18}$$

Then, for all $i \in \mathcal{P}_l$ (see Definition 1), the individual cost $J_i(\Phi_r, \theta_0)$ in the stealth configuration is strictly smaller than the one in the non-stealth configuration. Recall also that, for the nodes i with $i \neq l$ and $i \notin \mathcal{P}_l$, $J_i(\Phi_r, \theta_0) = 0$ in both configurations.

Proof. See Appendix A. ∎

The advantage of the stealth configuration will be further discussed in Section 4.

3.3 Identification of one given module

Let us now show how we can derive an estimate of $\theta_{0,l}$ using a model structure $\mathcal{M}_l = \{G_l(z,\theta_l), H_l(z,\theta_l) \mid \theta_l \in \mathbf{R}^{n_{\theta_l}}\}$ for the node $\mathcal{S}_l$ and a data set $Z_l^N = \{y_l(t), u_l(t) | t = 1,\dots,N\}$ collected as shown in Figure 3 in the stealth or the non-stealth setting. Let us first make the following assumptions:

Assumption 3 *We suppose that $\mathcal{M}_l$ is a full order model structure that is globally identifiable at $\theta_{0,l}$ i.e., $\theta_l = \theta_{0,l}$ is the only parameter vector for which $G_l(z,\theta_l)$ and $H_l(z,\theta_l)$ corresponds to $\mathcal{S}_l$. We also suppose that $K_l(z)G_l(z,\theta_{0,l})$ contains at least one delay, that $ref_{ext}(t)$ is a stationary signal and finally that $ref_{ext}(t)$, $r(t)$ and $\bar{e}$ are mutually uncorrelated.*

Then, using Z_l^N and $\mathcal{M}_l$, an estimate $\hat{\theta}_{N,l}$ of $\theta_{0,l}$ can be obtained via single-input single-output (SISO) prediction error identification [13]:

$$\hat{\theta}_{N,l} = \arg\min_{\theta_l} \frac{1}{N}\sum_{t=1}^{N} \epsilon_l^2(t,\theta_l) \tag{19}$$

with $\epsilon_l(t,\theta_l) = H_l^{-1}(z,\theta_l)\left(y_l(t) - G_l(z,\theta_l)u_l(t)\right)$. The consistency of (19) can be proven using a similar reasoning as in Theorem 1 of [1]. Since the noises e_i of nodes i having a path to Node l contribute to the excitation of $\mathcal{S}_l$ through the signal u_l used in (19) (see (10)), $\hat{\theta}_{N,l}$ is consistent even if $r = ref_{ext} = 0$ when Assumption 3 holds. If there is no path from any other node to Node l, $\hat{\theta}_{N,l}$ is also consistent under Assumption 3, provided a simple condition on the support of the spectrum $\Phi_r(\omega)$ of r is also respected (see [1,14] for more details). Note also that the derivation of (19) does not require any information on the order and the value of the plant transfer functions $G_i(z,\theta_{0,i})$ and $H_i(z,\theta_{0,i})$ for $i \neq l$.

Furthermore, $\hat{\theta}_{N,l}$ is also (asymptotically) normally distributed around $\theta_{0,l}$ with a covariance matrix P_{θ_l} that can be estimated using Z_l^N and $\hat{\theta}_{N,l}$ and whose inverse has the following expression [13]:

$$\begin{aligned} &P_{\theta_l}^{-1}(\Phi_r,\theta_0) = M(\theta_0) + \dots \\ &\dots + \frac{N}{2\pi\sigma_{e_l}^2}\int_{-\pi}^{\pi} M_F(e^{j\omega},\theta_{0,l})|R_l^u(e^{j\omega},\theta_0)|^2\Phi_r(\omega)d\omega \end{aligned} \tag{20}$$

where R_l^u is the l^{th} entry of R^u, $\sigma_{e_l}^2$ is the variance of e_l, $M_F(e^{j\omega},\theta_l) = F_l(e^{j\omega},\theta_l)F_l^*(e^{j\omega},\theta_l)$ ($F_l(z,\theta_l) = H_l^{-1}(z,\theta_l)\frac{\partial G_l(z,\theta_l)}{\partial\theta_l}$) and $M(\theta_0)$ is the joint contribution of $\bar{e}$ and ref_{ext} to the accuracy of the estimate (see Appendix B). We observe that $P_{\theta_l}^{-1}(\Phi_r,\theta_0)$ is an affine function of the power spectrum Φ_r of the excitation signal r (and a more complex function of θ_0). Equation (20) pertains to the stealth configuration. In the non-stealth configuration, we can use the same expression for $P_{\theta_l}(\Phi_r,\theta_0)$, but we have to replace R_l^u by $R_l^{u,NS}$ (see (16)).

The inverse of $P_{\theta_l}(\Phi_r,\theta_0)$ is a measure of the accuracy of the estimate $\hat{\theta}_{N,l}$ of $\theta_{0,l}$. In this paper, we will suppose that this accuracy will be deemed satisfactory if the following accuracy constraint $P_{\theta_l}^{-1}(\Phi_r,\theta_0) \geq R_{adm}$ is satisfied. The matrix R_{adm} is a given strictly positive-definite and symmetric matrix that reflects the desired accuracy.

Remark. For (20) to be a measure of the accuracy of $\hat{\theta}_{N,l}$, it is crucial that $\hat{\theta}_{N,l}$ be identified in a full-order model structure $\mathcal{M}_l$ (see Assumption 3). The latter can seem a strong assumption. Note however that there exist techniques based on the analysis of the prediction error to build such a full order model structure (see Section 16.6 in [13]).

4 Optimal experiment design problem

As mentioned in the introduction, we will design the spectrum Φ_r of the excitation signal r of the identification experiment described in the previous section in such a way that the accuracy constraint $P_{\theta_l}^{-1}(\Phi_r,\theta_0) \geq R_{adm}$ is satisfied with the smallest cost $J(\Phi_r,\theta_0)$ (see (14)). This optimization problem can thus be formulated as follows:

$$\min_{\Phi_r} J(\Phi_r,\theta_0) \quad \text{subject to} \tag{21}$$

$$P_{\theta_l}^{-1}(\Phi_r,\theta_0) \geq R_{adm} \tag{22}$$

This optimization problem can be considered both in the stealth and in the non-stealth setting by using the respective expressions for $J(\Phi_r,\theta_0)$ and for $P_{\theta_l}^{-1}(\Phi_r,\theta_0)$ in both cases. Before discussing how this optimization problem can be solved in practice, let us formulate the following result that illustrates the advantage of the stealth configuration. For this purpose, let us here also exclude some pathological cases:

Assumption 4 *Consider the notations introduced in Section 3 for the identification of Node l and in particular the l^{th} entries $R_l^u(z,\theta_0)$ and $R_l^{u,NS}(z,\theta_0)$ of the vectors of transfer functions $R^u(z,\theta_0)$ and $R^{u,NS}(z,\theta_0)$ defined in (11) and in (16), respectively. We assume that $R_l^u(e^{j\omega},\theta_0)$ and $R_l^{u,NS}(e^{j\omega},\theta_0)$ are equal to zero only at those frequencies ω where the frequency response $K_l(e^{j\omega})$ of the controller present in Node l is infinite*[1] *(due to an integrator or a resonator).*

Proposition 2 *Consider an identification experiment in Node l of a network satisfying Assumptions 1, 2, and 4. Consider, for this identification experiment, the optimal experiment design problem (21)-(22) in the stealth and in the non-stealth setting (i.e. using the respective expressions for $J(\Phi_r,\theta_0)$ and for $P_{\theta_l}^{-1}(\Phi_r,\theta_0)$ in both cases) and let us denote by $\Phi_{r,opt}^S$ and $\Phi_{r,opt}^{NS}$ the optimal spectra obtained in these two settings. Then, we have that the optimal cost $J(\Phi_{r,opt}^S,\theta_0)$ in the stealth setting is strictly*

[1] At those frequencies ω where $K_l(e^{j\omega})$ is infinite, we have indeed that $S_{0,l}(e^{j\omega}) = T_{0,l}(e^{j\omega}) = T_{init,l}(e^{j\omega}) = 0$.

smaller than the cost $J(\Phi_{r,opt}^{NS}, \theta_0)$ in the non-stealth setting if the model $T_{init,l}$ of $T_{0,l}$ used in the stealth compensation (6) *has the following property for all ω where $\Phi_{r,opt}^{NS}(\omega) \neq 0$:*

$$\frac{|T_{0,l}(e^{j\omega}) - T_{init,l}(e^{j\omega})|}{|T_{0,l}(e^{j\omega})|} < \frac{|R_l^u(e^{j\omega}, \theta_0)|}{|R_l^{u,NS}(e^{j\omega}, \theta_0)|} \qquad (23)$$

where R_l^u (resp. $R_l^{u,NS}$) is the l^{th} entry of R^u (resp. $R^{u,NS}$) defined in (11) *(resp.* (16)*).*
Proof. See Appendix C. ∎

As shown in Proposition 2, we thus see that the stealth configuration, which can be very easily implemented in a multi-agent network, will be, in many cases, advantageous to obtain the required accuracy for the model of $\mathcal{S}_l$ with the smallest possible identification cost. The condition (23) on $T_{init,l}$ is more complex than (18). However, (23) will be respected if $T_{init,l}$ is not a too poor estimate of $T_{0,l}$.

Let us now turn to the problem of solving the optimal experiment design problem (21)-(22). If θ_0 would be known, this optimization problem could be solved using convex optimization since $J(\Phi_r, \theta_0)$ and $P_{\theta_l}^{-1}(\Phi_r, \theta_0)$ are affine function of the to-be-designed spectrum $\Phi_r(\omega)$. Different spectrum parametrizations can be considered for this purpose: the moment approach [2] [19,17], finite dimensional expansions of the spectrum resulting in filtered white noise [9,2] or sum-of-sinusoids (multisine) solutions [9,2]. Note that the last two parametrizations restrict the class of spectra in which the optimum spectrum is determined, whereas the moment approach cannot handle frequency-wise constraints.

Let us now deal with the fact that θ_0 is unknown, i.e. the classical chicken-and-egg issue in optimal experiment design.

5 Dealing with the chicken-and-egg issue

5.1 Approach based on $\theta_{init,l}$

A commonly used approach to circumvent this so-called chicked-and-egg problem is to replace θ_0 by an initial estimate $\theta_{init} = (\theta_{init,1}^T, \ldots, \theta_{init,N_{mod}}^T)^T$. In this paper, we have until now assumed (see Section 3.1) that we have an initial estimate $\theta_{init,l}$ of the to-be-identified parameter vector $\theta_{0,l}$ (and thus not of the whole parameter vector θ_0). This initial estimate $\theta_{init,l}$ (e.g., obtained via a previous identification experiment in a full order model structure $\mathcal{M}_l$ for $\mathcal{S}_l$) will be crucial to deal with the chicken-and-egg issue of (21)-(22).

Let us first consider the case of an homogenous network. In this case, the initial estimate $\theta_{init,l}$ for $\theta_{0,l}$ is also an initial estimate for $\theta_{0,i}$ $(i \neq l)$ and thus we can use $\theta_{init} = (\theta_{init,l}^T, \ldots, \theta_{init,l}^T)^T$ as initial estimate for θ_0. Consequently, the optimal spectrum can thus be determined by minimizing $J(\Phi_r, \theta_{init})$ subject to $P_{\theta_l}^{-1}(\Phi_r, \theta_{init}) \geq R_{adm}$. In the expression of $P_{\theta_l}^{-1}(\Phi_r, \theta_{init})$ (see Appendix B), Σ is approximated by the sample covariance of $\bar{H}^{-1}(z, \theta_{init}) \left(\bar{y} - \bar{G}(z, \theta_{init})\bar{u}\right)$ computed e.g. with normal operation data. The above optimization problem can be solved in the non-stealth setting and in the stealth setting. In the stealth setting, since $\theta_{init,l}$ is also used to compute $x(t)$ (see (6)), the cost $J(\Phi_r, \theta_{init})$ reduces to $J_l(\Phi_r, \theta_{init})$.

In the case of a non-homogenous network, the initial estimate $\theta_{init,l}$ for $\theta_{0,l}$ is not an initial estimate for $\theta_{0,i}$ $(i \neq l)$, but the stealth configuration will nevertheless allow one to deal with the chicken-and-egg issue even if we do not have initial estimates for $\theta_{0,i}$ $(i \neq l)$. To show this, let us replace $\theta_{0,l}$ by $\theta_{init,l}$ in the expression for $R^u(z, \theta_0)$ (see (11)). When we do this, $R^u(z, \theta_0)$ reduces to $\bar{m}_l S_{init,l}(z)$ with $S_{init,l}(z) = (1 + K_l(z)G_l(z, \theta_{init,l}))^{-1}$ and the cost J reduces to the individual cost J_l (see (15)). Consequently, if, in the expression (20) for P_θ^{-1}, we neglect [3] the positive-definite term $M(\theta_0)$ which cannot be approximated with only $\theta_{init,l}$, the optimal experiment design problem can be tackled using the following optimization problem which does not require initial estimates of $\theta_{0,i}$ for $i \neq l$:

$$\min_{\Phi_r} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(|G_l(e^{j\omega}, \theta_{init,l})|^2 + \eta\right) |S_{init,l}(e^{j\omega})|^2 \Phi_r(\omega)\, d\omega$$
$$\text{s.t. } \frac{N}{2\pi\sigma_{e_l}^2} \int_{-\pi}^{\pi} M_F(e^{j\omega}, \theta_{init,l}) |S_{init,l}(e^{j\omega})|^2 \Phi_r(\omega) d\omega \geq R_{adm}$$

where $\sigma_{e_l}^2$ can be approximated by the sample variance of $H_l^{-1}(z, \theta_{init,l}) (y_l - G_l(z, \theta_{init,l})u_l)$ computed e.g. with normal operation data. It is to be noted that this optimization problem can only be considered in the stealth stetting (unlike $R^u(z, \theta_0)$, $R^{u,NS}(z, \theta_0)$ does indeed not reduce to $\bar{m}_l S_{init,l}(z)$ when we replace $\theta_{0,l}$ by $\theta_{init,l}$).

Note that the three spectrum parametrizations discussed at the end of Section 4 can be used to tackle the optimization problems presented in this subsection.

5.2 Robust approach

As we will see in the sequel, if we have more information about the network, we will be able to deal with the chicken-and-egg issue in a more robust way. This will be possible if we have a parametric uncertainty region U_{init} containing the true parameter vector θ_0. Such information about the network can e.g., be available if the different systems $\mathcal{S}_i$ $(i = 1, ..., N_{mod})$ that make up the network have previously been identified in full order model structures $\mathcal{M}_i$ using the prediction error method of Section 3.3. In this case, the corresponding identified parameter vectors $\theta_{init,i}$ are all normally distributed around $\theta_{0,i}$ $(i = 1, ..., N_{mod})$ and an ellipsoidal uncertainty region $U_{init,i}$ containing $\theta_{0,i}$ modulo

[2] which is also known as the partial correlation expansion approach [9]

[3] An alternative is to estimate $M(\theta_0)$ using an identification of $\mathcal{S}_l$ with normal operation data (see [15]).

a user-chosen probability level can be built based on the covariance matrix $P_{init,i}$ of $\theta_{init,i}$ i.e., $U_{init,i} := \left\{\theta_i \in \mathbf{R}^{n_{\theta_i}} \mid (\theta_i - \theta_{init,i})^T P_{init,i}^{-1}(\theta_i - \theta_{init,i}) < \chi\right\}$ where χ depends on the chosen probability level [13]. The ellipsoidal uncertainty regions $U_{init,i}$ $(i = 1, ..., N_{mod})$ obtained in these successive identification experiments can be regrouped to yield $U_{init} = \{\theta = (\theta_1^T, ..., \theta_{N_{mod}}^T)^T \mid \theta_i \in U_{init,i}\ (i = 1, ..., N_{mod})\}$. In the case of an homogenous network, the uncertainty region U_{init} can obviously be obtained based on the initial identification of one single module.

If the uncertainty set U_{init} is available, we can then robustify our optimal experiment design as follows:

$$\min_{\Phi_r,\gamma} \gamma \quad \text{subject to} \tag{24}$$

$$J(\Phi_r, \theta) \leq \gamma \quad \forall \theta \in U_{init} \quad \text{and} \tag{25}$$

$$P_{\theta_l}^{-1}(\Phi_r, \theta) \geq R_{adm} \quad \forall \theta \in U_{init} \tag{26}$$

Besides the classical advantages of a robustified optimal experiment design problem [16], the cost J will, in the above formulation, no longer reduce to J_l in the stealth setting (since we do no longer replace $\theta_{0,l}$ by $\theta_{init,l}$ in the expression of the cost). Consequently, the robustified formulation will favour spectra Φ_r yielding, for all $\theta \in U_{init}$, small perturbations $R_i^y(z,\theta)r(t)$ and $R_i^u(z,\theta)r(t)$ not only for $i = l$, but also for all $i \in \mathcal{P}_l$ (see (9)-(10)). For nodes $i \in \mathcal{P}_l$, this e.g. means that the power of $R_i^u(z,\theta)r(t) = \mathcal{N}_i(z,\theta)\left(\frac{G_l(z,\theta)}{1+K_l(z)G_l(z,\theta)} - T_{init,l}(z)\right)r(t)$ has to be made small for all $\theta \in U_{init}$. Consequently, the robustified optimal experiment design problem will generally and among other considerations favour spectra $\Phi_r(\omega)$ with more contributions in the frequency ranges where the stealth compensation will be more effective due to a small uncertainty of $T_{init,l}(z)$ (thereby *robustifying* the stealth configuration).

Finding a tractable approach to deal with the robustified optimal experiment design problem (24)-(26) is a complex issue. Given the complexity of the network, the general approach would be to replace the initial uncertainty set (containing an infinite number of elements) by a number n_g of grid points of this uncertainty set U_{init} (see e.g. [9,2,16]). In this approach, the cost constraint (25) and the accuracy constraint (26) are replaced each by n_g constraints (one for each grid point). This so-called gridding approach of course entails an approximation since ensuring these n_g constraints does not imply ensuring (25) and (26).

In this paper, we will use this gridding approach for the accuracy constraint i.e., (26) is replaced by the n_g constraints $P_{\theta_l}^{-1}(\Phi_r, \theta^j) \geq R_{adm}$ $(j = 1, ..., n_g)$ where $\theta^j \in U_{init}$ $(j = 1, ..., n_g)$.

For the cost constraint (25), we will propose a better approach than the gridding approach. Indeed, we will replace (25) by an alternative constraint that is linear in γ and Φ_r and that implies (25). In order to derive such a tractable constraint implying (25), we will need to restrict attention to a multisine parametrization for the to-be-determined spectrum (see the end of Section 4):

$$\Phi_r(\omega) = \pi \sum_{m=1}^{L} c_m \ (\delta(\omega - \omega_m) + \delta(\omega + \omega_m)) \geq 0 \ \forall \omega \tag{27}$$

where $c_m \geq 0$ $(m = 1, \ldots, L)$ will be the decision variables of the optimization problem and where the frequencies ω_m $(m = 1, \ldots, L)$ are fixed by the user (as e.g. a fine grid of the frequency range $[0\ \pi]$). We have then the following result:

Proposition 3 *Consider the robust cost constraint* (25) *and a spectrum of the type* (27)*. Then, the constraint* (25) *holds for a given* γ *if the following inequality linear in the decision variables* c_m $(m = 1, \ldots, L)$ *holds:*

$$\sum_{m=1}^{L} c_m \ \alpha(\omega_m) \leq \gamma \tag{28}$$

where $\alpha(\omega)$ $(m = 1, ..., L)$ *is an upper bound for* $\mathcal{J}(\omega) = \sup_{\theta \in U_{init}} \left(\mathcal{R}_\eta^*(e^{j\omega}, \theta)\mathcal{R}_\eta(e^{j\omega}, \theta)\right)$

Proof. Using (27) and (14), the robust cost constraint (25) can be rewritten as follows:

$$\sum_{m=1}^{L} c_m \ \left(\mathcal{R}_\eta^*(e^{j\omega_m}, \theta)\mathcal{R}_\eta(e^{j\omega_m}, \theta)\right) \leq \gamma \qquad \forall \theta \in U_{init} \tag{29}$$

Since, for any frequency ω_m, we have that $\left(\mathcal{R}_\eta^*(e^{j\omega_m}, \theta)\mathcal{R}_\eta(e^{j\omega_m}, \theta)\right) \leq \alpha(\omega_m)\ \forall \theta \in U_{init}$, it is clear that (29) holds if (28) holds. ■

Let us observe that, at each ω, an upper bound $\alpha(\omega)$ for $\mathcal{J}(\omega)$ can be computed by following the hierarchical approach introduced in [1] (see [14] for more details). Let us also observe that (28) is an inequality constraint linear in the coefficients c_m $(m = 1, \ldots, L)$. Moreover, since Φ_r is affine in c_m $(m = 1, \ldots, L)$, the n_g accuracy constraints replacing (26) are also linear in these coefficients. Consequently, the optimization problem (24)-(26) can be tackled by the LMI optimization problem consisting of determining the coefficients $c_m \geq 0$ $(m = 1, \ldots, L)$ minimizing the value of γ subject to (28) and to $P_{\theta_l}^{-1}(\Phi_r, \theta^j) \geq R_{adm}$ $(j = 1, ..., n_g)$.

Let us denote by $c_{m,opt}$ $(m = 1, \ldots, L)$ and by γ_{opt} the solution of this LMI optimization problem and let us also denote by $\Phi_{r,opt}$ the spectrum corresponding to the coefficients $c_{m,opt}$. Then, due to Proposition 3, we have that γ_{opt} is an upper bound of $\sup_{\theta \in U_{init}} J(\Phi_{r,opt}, \theta)$ and thus of $J(\Phi_{r,opt}, \theta_0)$.

This property is a clear advantage of the proposed approach to robustify the cost constraint with respect to the gridding approach. Note however that (28) implies, but is not equivalent to the original constraint (25)/(29). Consequently, using (28) instead of (25) comes with a dose of conservatism. We have first to restrict attention to (27) and there are also two sources of conservatism

when we replace (29) by (28). The first source consists in going from (29) to $\sum_{m=1}^{L} c_m \mathcal{J}(\omega_m) \leq \gamma$ and the second one in going from $\sum_{m=1}^{L} c_m \mathcal{J}(\omega_m) \leq \gamma$ to (28). Let us explain these sources of conservatism in more details. In the first source of conservatism, we consider the optimization problem leading to $\mathcal{J}(\omega_m)$ independently at each ω_m ($m = 1, ..., L$). Consequently, the parameter θ at which the supremum $\mathcal{J}(\omega_m)$ is obtained can be different at each frequency ω_m (unlike in the original constraint (29)). In the second source of conservatism, we replace the quantity $\mathcal{J}(\omega_m)$ by an upper bound $\alpha(\omega_m)$ computed via the hierarchical approach.

6 Numerical illustration

Let us illustrate our results with a numerical example. We consider the network of Figure 1 made up of $N_{mod} = 6$ homogenous nodes. In other words, the true systems $\mathcal{S}_i$ ($i = 1, ..., 6$) are all identical and given by the following ARX system [12] with two resonance peaks: $y_i(t) = (z^{-3}B_0(z))/(A_0(z))u_i(t)+(1)/(A_0(z))e_i(t)$ with $B_0(z) = 0.10276+0.18123z^{-1}$, $A_0(z) = 1-1.99185z^{-1}+2.20265z^{-2} - 1.84083z^{-3} + 0.89413z^{-4}$. The variances of the white noises e_i will also all be chosen equal to 0.5 and $\Sigma = 0.5I_{N_{mod}}$. We further suppose that these true systems are all controlled by the same local controller $K(z) = \frac{K_B(z)}{K_A(z)}$ with $K_B(z) = 0.03742 - 0.06719z^{-1} + 0.06995z^{-2} - 0.03814z^{-3} - 0.02546z^{-4} + 0.06323z^{-5} - 0.04707z^{-6} + 0.03222z^{-7}$ and $K_A(z) = 1 - 3.348z^{-1} + 5.953z^{-2} - 7.163z^{-3} + 6.143z^{-4} - 3.705z^{-5} + 1.368z^{-6} - 0.2482z^{-7}$.

For simplicity, the initial information needed to deal with the chicken-and-egg issue is determined from one (cheap) open-loop experiment on the ARX system (disconnected from the network) with a white input signal of variance 1.9 and of duration $N = 1000$. This open-loop experiment yields the following identified parameter vector $\hat{\theta}_{mod} = (-1.9949, 2.1871, -1.8042, 0.8764, 0.0807, 0.1671)^T$ of the true parameter vector $\theta_{0,mod}$ of the ARX system. Using the covariance matrix P_{mod} of $\hat{\theta}_{mod}$, we derive the following uncertainty set $U_{mod} := \left\{\theta_{mod} \mid (\theta_{mod} - \hat{\theta}_{mod})^T P_{mod}^{-1} (\theta_{mod} - \hat{\theta}_{mod}) < 12.6\right\}$ that is guaranteed to contain $\theta_{0,mod}$ with a probability level of 95% and that here effectively contains $\theta_{0,mod}$. Since the network is homogenous, an initial estimate for $\theta_0 = (\theta_{0,mod}^T, \theta_{0,mod}^T, ..., \theta_{0,mod}^T)^T \in \mathbf{R}^{36}$ is given by $\theta_{init} = (\hat{\theta}_{mod}^T, \hat{\theta}_{mod}^T, ..., \hat{\theta}_{mod}^T)^T$ and the uncertainty set U_{init} has the expression given in Section 5.2 with $U_{init,i} = U_{mod}$ ($i = 1, ..., 6$)

Our objective will be to design the spectrum $\Phi_r(\omega)$ of the excitation signal r that has to be added to Node 5 (i.e. $l = 5$) during an identification experiment of duration $N = 1000$ to improve the accuracy of the model of $\mathcal{S}_5$ in such a way that the following accuracy constraint is satisfied $P_{\theta_5}^{-1}(\Phi_r, \theta_0) > R_{adm}$ where R_{adm} is chosen as the inverse of the diagonal matrix $(diag(0.0199,\ 0.022,\ 0.018,\ 0.009,\ 0.005,\ 0.009))^2$. We will furthermore suppose that $ref_{ext}(t) = 0$ during the identification experiment. In addition, we define the cost of the identification experiment as in (14) with $\eta = 1$ and we parametrize $\Phi_r(\omega)$ as in (27) with $L = 20$ frequencies distributed in the frequency range $\left[10^{-2},\ \pi\right]$. In the stealth setting, the stealth compensation is implemented using (6) with the transfer function $T_{init,5}$ that can be constructed based on $\theta_{init,5} = \hat{\theta}_{mod}$. Moreover, for the robustification of the accuracy constraint, we consider 20 grid points in U_{init} (one of these 20 grid points is θ_{init}).

Let first consider the robustified approach of Section 5.2 in the stealth setting and let us solve the LMI optimization given at the end of this section. This yields $\gamma_{opt}^S = 179.92$ and an optimal spectrum $\Phi_{r,opt}^S$ having only contributions at three frequencies i.e., $\omega = 0.45$, $\omega = 1.35$ and $\omega = 1.65$. As discussed in Section 5.2, the value γ_{opt}^S is an upper bound for $\sup_{\theta \in U_{init}} J(\Phi_{r,opt}^S, \theta)$ and thus also an upper bound for the a-priori unknown cost $J(\Phi_{r,opt}^S, \theta_0)$ which is here equal to 130.22. The individual costs $J_i(\Phi_{r,opt}^S, \theta_0)$ for $i = 5$ and for all $i \in \mathcal{P}_5 = \{2, 3, 4, 6\}$ are given in Table 1.

Let us compare these results with what is obtained in the non-stealth setting. Let us for this purpose solve the LMI optimization given at the end of Section 5.2 for this setting. This yields $\gamma_{opt}^{NS} = 228.35$ and the optimal spectrum $\Phi_{r,opt}^{NS}$ has only contributions at four frequencies i.e., $\omega = 0.45$, $\omega = 1.35$, $\omega = 1.65$ and $\omega = 3$. The actual cost $J(\Phi_{r,opt}^{NS}, \theta_0)$ of an experiment with $\Phi_{r,opt}^{NS}$ is equal to 151.29. This is 15% larger than in the stealth case where $J(\Phi_{r,opt}^S, \theta_0) = 130.22$. Moreover, as shown in Table 1, the individual costs J_i for $i \in \mathcal{P}_5$ are approximatively 100 times larger than when the stealth configuration is implemented. This clearly shows the advantage of the stealth configuration.

As mentioned in Section 5.2, we only partially robustify the accuracy constraint since we replace the infinite set U_{init} by 20 grid points. Despite this approximation, the constraint (22) is respected both in the stealth setting and in the non-stealth setting (i.e. with $\Phi_{r,opt}^S$ and $\Phi_{r,opt}^{NS}$). The robustification of the cost constraint is not based on an approximation, but entails a certain conservatism. Let us thus check the extent of this conservatism; e.g. in the stealth setting. For this purpose, we compute a lower bound J^{lb} for $\sup_{\theta \in U_{init}} J(\Phi_{r,opt}^S, \theta)$ by considering the maximal value of $J(\Phi_{r,opt}^S, \theta^i)$ over a set of 1000 grid points $\theta^i \in U_{init}$. This procedure yields $J^{lb} = 167.60$. Consequently, the conservatism of our procedure remains limited in this example since there is only 7% of difference between J^{lb} and the upper bound J^{ub} for this quantity obtained via our approach i.e., $J^{ub} = \gamma_{opt}^S = 179.92$.

Let us now consider the non-robustified approach of Section 5.1 and let us focus on the stealth setting for the sake of brevity. Since the network is homogenous, we replace the whole vector θ_0 by θ_{init} and we determine $\Phi_{r,opt}^{nr,S}$ by minimizing $J(\Phi_r, \theta_{init})$ subject to $P_{\theta_l}^{-1}(\Phi_r, \theta_{init}) \geq$

R_{adm}. The optimal spectrum $\Phi_{r,opt}^{nr,S}$ has only contributions at the same three frequencies as $\Phi_{r,opt}^{S}$. While the constraint (22) is respected with $\Phi_{r,opt}^{nr,S}$, the predicted cost $J(\Phi_{r,opt}^{nr,S}, \theta_{init}) = 119.22$ underestimates the actual cost $J(\Phi_{r,opt}^{nr,S}, \theta_0) = 126.32$. The non-robustified approach does indeed not give the same guarantee as the robustified one. In Table 1, we also observe that, for $i \in \mathcal{P}_5 = \{2, 3, 4, 6\}$, the individual costs $J_i(\Phi_{r,opt}^{S}, \theta_0)$ in the robustified case are all 25% smaller than the individual costs $J_i(\Phi_{r,opt}^{nr,S}, \theta_0)$ in the non-robustified case. This shows the advantage of the achieved robustification of the stealth configuration discussed in Section 5.2. Note finally that the total cost $J(\Phi_{r,opt}^{nr,S}, \theta_0) = 126.32$ obtained with this non-robustified approach is slightly smaller that the total cost $J(\Phi_{r,opt}^{S}, \theta_0) = 130.22$ obtained in the robustified version. Since $\Phi_{r,opt}^{S}$ has to satisfy the accuracy and the cost constraints for more θ than just θ_{init}, this spectrum will indeed generally be larger than $\Phi_{r,opt}^{nr,S}$.

Table 1
Individual costs J_i $(i = 2, ..., 6)$ obtained using the optimal spectra $\Phi_{r,opt}^{S}$ and $\Phi_{r,opt}^{NS}$ $\Phi_{r,opt}^{nr,S}$

	J_2	J_3	J_4	J_5	J_6
$\Phi_{r,opt}^{S}$	0.0023	0.0178	0.0178	130.0957	0.0853
$\Phi_{r,opt}^{NS}$	0.2197	1.7196	1.7196	139.3723	8.2553
$\Phi_{r,opt}^{nr,S}$	0.0029	0.0230	0.0230	126.1607	0.1104

7 Alternative to the stealth configuration

As mentioned in Section 3, the stealth configuration allows to reduce the propagation of $r(t)$ to the other nodes of the network while maintaining the tracking performance and the disturbance rejection performance of the network intact. This is achieved by choosing, for Node l, the signal $\tilde{y}_l(t)$ (i.e., the signal that is transmitted to compute $\bar{y}_{ref}(t)$) as $\tilde{y}_l(t) = y_l(t) - x(t)$ with $x(t)$ as given in (6). A valid alternative to the stealth approach would be an approach that entirely cancels the propagation of $r(t)$ to the other nodes and that would maintain a satisfactory tracking and disturbance rejection performance. One could attempt to achieve this objective by choosing $\tilde{y}_l(t) = y_{ref,l}(t)$ or, as proposed by anonymous reviewers, $\tilde{y}_l(t) = K_l(z)T_{init,l}(z)y_{ref,l}(t)$. With these choices for $\tilde{y}_l(t)$, the direct advantage is to entirely cancel the propagation of $r(t)$ (and of $v_l(t)$) to the other nodes while only slightly changing the conditions for the experiment design in Node l ($R_l^u(z, \theta_0) = S_{0,l}(z)$ in these configurations, but that is also approximately the case in the stealth configuration). It is also clear that these choices for $\tilde{y}_l(t)$ modifies the transfer vectors R_{ext}^{y} and R_{ext}^{u} (measuring the tracking performance of the network) and the transfer matrices S^y and S^u (measuring the disturbance rejection performance of the network). Consequently, what is gained by canceling the propagation of $r(t)$ to the other nodes could be easily lost via an eventual stronger decrease of the tracking performance or via a control input becoming excessive (remember indeed that the stealth configuration allows to strongly reduce this propagation of $r(t)$ to the other nodes). Note also that these alternative choices for $\tilde{y}_l(t)$ may even lead to the instability of the network in some unfortunate cases. This said, if a robustness analysis can be performed to analyze the modification of the tracking and disturbance rejection performance, it is certainly useful to compare these approaches with the stealth approach to determine which one leads to the smaller perturbation of the network performance during the identification experiment. However, the a-priori information about the unknown dynamics (i.e. about the unknown θ_0) may be insufficient to perform this robustness analysis. This is clearly the case when the only a-priori information we have on the network is an initial estimate $\theta_{init,l}$ of $\theta_{0,l}$ (see Section 5.1). In this case, it is our opinion that the stealth approach has to be preferred due to its guarantee of maintaining the network stability and its guarantee of maintaining the tracking performance and the disturbance rejection performance of the network intact.

8 Conclusions

This paper extends the least costly identification experiment design framework to the case of the identification of one module in a network of locally controlled systems. The cost of the identification experiment (that is minimized under a certain accuracy constraint) is here defined as a function of the perturbations induced by the excitation signal on the input and output signals of each module. The propagation of the influence of the excitation signal can be further reduced by an extension of the stealth identification paradigm or using the alternatives presented in Section 7.

References

[1] X. Bombois, A. Korniienko, H. Hjalmarsson, and G. Scorletti. Optimal identification experiment design for the interconnection of locally controlled systems. *Automatica*, 89:169–179, 2018.

[2] X. Bombois, G. Scorletti, M. Gevers, P.M.J. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, 2006.

[3] A. Dankers, P. Van den Hof, X. Bombois, and P. Heuberger. Identification of dynamic models in complex networks with prediction error methods - predictor input selection. *IEEE Transactions on Automatic Control*, 61(4):937–952, 2016.

[4] N. Everitt, G. Bottegal, C. Rojas, and H. Hjalmarsson. On the variance analysis of identified linear MIMO models. In *Proc. 54th IEEE Conference on Decision and Control*, pages 1447–1452, Osaka, Japan, 2015.

[5] J.A. Fax and R.H. Murray. Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49(9):1465–1476, 2004.

[6] M. Gevers and A.S. Bazanella. Identification in dynamic networks: identifiability and experiment design issues. In *Proc. 54th IEEE Conference on Decision and Control*, pages 4005–4010, Osaka, Japan, 2015.

[7] A. Haber and M. Verhaegen. Moving horizon estimation for large-scale interconnected systems. *IEEE Transactions on Automatic Control*, 58(11):2834–2847, 2013.

[8] P. Hägg and B. Wahlberg. On identification of parallel cascade serial systems. In *Proc. IFAC World Congress*, pages 9978–9983, Cape Town, South Africa, 2014.

[9] H. Jansson and H. Hjalmarsson. Input design via LMIs admitting frequency-wise model specifications in confidence regions. *IEEE Transactions on Automatic Control*, 50(10):1534–1549, October 2005.

[10] A. Korniienko, X. Bombois, H. Hjalmarsson, and G. Scorletti. Hierarchical robust analysis for identified systems in network. *IFAC-PapersOnLine*, 51(25):383–389, 2018.

[11] A. Korniienko, G. Scorletti, E. Colinet, and E. Blanco. Performance control for interconnection of identical systems: Application to PLL network design. *International Journal of Robust and Nonlinear Control*, 2014.

[12] I. Landau, D. Rey, A. Karimi, A. Voda, and A. Franco. A flexible transmission system as a benchmark for robust digital control. *European Journal of Control*, 1(2):77–96, 1995.

[13] L. Ljung. *System Identification: Theory for the User, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ, 1999.

[14] F. Morelli, X. Bombois, H. Hjalmarsson, L. Bako, and K. Colin. Least costly identification experiment for the identification of one module in a dynamic network. Technical report, 2019. Available on the HAL repository under the reference hal-02387776.

[15] M. Potters, X. Bombois, M. Forgione, P. Modén, M. Lundh, H. Hjalmarsson, and P. Van den Hof. Optimal experiment design in closed loop with unknown, nonlinear and implicit controllers using stealth identification. In *European Control Conference (ECC)*, pages 726–731, 2014.

[16] C. Rojas, J. Welsh, G. Goodwin, and A. Feuer. Robust optimal experiment design for system identification. *Automatica*, 43:993–1008, 2007.

[17] P. Stoica and T. Söderström. A Useful Parameterization for Optimal Experiment Design. *IEEE Transactions on Automatic Control*, AC-27(4):986–989, August 1982.

[18] H. Weerts, A. Dankers, and P. Van den Hof. Identifiability in dynamic network identification. In *Proc. IFAC Symposium on System Identification*, pages 1409–1414, Beijing, China, 2015.

[19] M. Zarrop. *Design for Dynamic System Identification*. Lecture Notes in Control and Inform. Sci. 21, Springer Verlag, Berlin, New-York, 1979.

A Proof of Proposition 1

Let us consider the expression (15) for the individual costs $J_i(\Phi_r, \theta_0)$ $(i = 1, ..., N_{mod})$ and introduce the following notations to distinguish them in the stealth (superscript S) and non-stealth case (superscript NS):

$$J_i^S(\Phi_r, \theta_0) = \dots \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(|G_i(e^{j\omega}, \theta_{0,i})|^2 + \eta\right) |R_i^u(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega)\, d\omega \tag{A.1}$$

$$J_i^{NS}(\Phi_r, \theta_0) = \dots \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(|G_i(e^{j\omega}, \theta_{0,i})|^2 + \eta\right) |R_i^{u,NS}(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega)\, d\omega \tag{A.2}$$

where R^u is given by (11) and $R^{u,NS}$ is given by (16). Let us now observe that, for $i \neq l$, we have that:

$$|R_i^u(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega) = \dots \\ \dots |\mathcal{N}_i(e^{j\omega}, \theta_0)|^2 |T_{0,l}(e^{j\omega}) - T_{init,l}(e^{j\omega})|^2 \Phi_r(\omega)$$

$$|R_i^{u,NS}(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega) = |\mathcal{N}_i(e^{j\omega}, \theta_0)|^2 |T_{0,l}(e^{j\omega})|^2 \Phi_r(\omega)$$

Let us first consider the indexes $i \in \mathcal{P}_l$ and recall Assumption 2. In this case, if $T_{init,l}$ satisfies (18) for all frequencies ω where $\Phi_r(\omega) \neq 0$, we have for all these frequencies that $|R_i^{u,NS}(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega) > |R_i^u(e^{j\omega}, \theta_0)|^2 \Phi_r(\omega)$. Consequently, for all $i \in \mathcal{P}_l$, we have that $J_i^S(\Phi_r, \theta_0) < J_i^{NS}(\Phi_r, \theta_0)$. Finally, for the indexes $i \neq l$, $i \notin \mathcal{P}_l$, $J_i^S(\Phi_r, \theta_0) = J_i^{NS}(\Phi_r, \theta_0) = 0$ since $\mathcal{N}_i(z, \theta_0) = 0$ in this case.

B Covariance matrix P_{θ_l} of $\hat{\theta}_{N,l}$

Since $\hat{\theta}_{N,l}$ is a consistent estimate and $\epsilon_l(t, \theta_{0,l}) = e_l(t)$, Chapter 9 of [13] shows that $\hat{\theta}_{N,l}$ is (asymptotically) normally distributed around $\theta_{0,l}$ and the inverse of the covariance matrix of $\hat{\theta}_{N,l}$ is given by $P_{\theta_l}^{-1} = \frac{N}{\sigma_{e_l}^2} \bar{E} \psi_l(t, \theta_{0,l}) \psi_l^T(t, \theta_{0,l})$ with $\psi_l(t, \theta_l) = \frac{-\partial \epsilon_l(t, \theta_l)}{\partial \theta_l}$ [13]. It is easy to show (see e.g. [2]) that $\psi_l(t, \theta_{0,l}) = F_l(z, \theta_{0,l}) u_l(t) + L_l(z, \theta_{0,l}) e_l(t)$ with $F_l(\theta_l)$ as defined below (20) and with $L_l(\theta_l) = H_l^{-1}(\theta_l) \frac{\partial H_l(\theta_l)}{\partial \theta_l}$. Replacing the expression (10) for u_l into the above expression for $\psi_l(t, \theta_{0,l})$ and recalling that r, ref_{ext} and $\bar{e}$ are mutually independent, we obtain the expression (20) with $M(\theta_0) = M_{\bar{e}}(\theta_0) + M_{ref_{ext}}(\theta_0)$:

$$M_{ref_{ext}} = \frac{N}{2\pi\sigma_{e_l}^2} \int_{-\pi}^{\pi} M_F(e^{j\omega}, \theta_{0,l}) |R_{ext,l}^u(e^{j\omega}, \theta_0)|^2 \Phi_{ref_{ext}}(\omega)\, d\omega$$

$$M_{\bar{e}} = \frac{N}{2\pi\sigma_{e_l}^2} \int_{-\pi}^{\pi} \mathcal{Z}_l(e^{j\omega}, \theta_0)\ \Sigma\ \mathcal{Z}_l^*(e^{j\omega}, \theta_0) d\omega$$

with $\Sigma = E\bar{e}(t)\bar{e}(t)^T$, $R_{ext,l}^u$ the l^{th} entry of R_{ext}^u, $\Phi_{ref_{ext}}$ the spectrum of ref_{ext}, and $\mathcal{Z}_l(z)$ a matrix of transfer functions of dimension $n_{\theta_l} \times N_{mod}$ whose l^{th} column is $L_l + F_l S_{ll}^u$ and whose k^{th} column $(k \neq l)$ is equal to $F_l S_{lk}^u$ (S_{lk}^u is the entry $l \times k$ of S^u in (10)).

C Proof of Proposition 2

Let us introduce the following notations to distinguish the covariance matrice $P_{\theta_l}(\Phi_r, \theta_0)$ in the stealth and non-stealth case and let us also consider the notations introduced in (A.1)-(A.2) to distinguish the individual costs in those two cases.

$$P_{\theta_l,S}^{-1}(\Phi_r^S, \theta_0) = M(\theta_0) + \dots \\ \dots + \frac{N}{2\pi\sigma_{e_l}^2} \int_{-\pi}^{\pi} M_F(e^{j\omega}, \theta_{0,l}) |R_l^u(e^{j\omega}, \theta_0)|^2 \Phi_r^S(\omega)\, d\omega$$

$$P^{-1}_{\theta_l,NS}(\Phi_r^{NS},\theta_0) = M(\theta_0) + \dots$$
$$\dots + \tfrac{N}{2\pi\sigma^2_{e_l}} \int_{-\pi}^{\pi} M_F(e^{j\omega},\theta_{0,l}) |R_l^{u,NS}(e^{j\omega},\theta_0)|^2 \Phi_r^{NS}(\omega)\, d\omega$$

Using the optimal spectrum $\Phi^{NS}_{r,opt}$ in the non-stealth case, let us define the following spectrum:

$$\Phi_r^S(\omega) = \frac{|R_l^{u,NS}(e^{j\omega},\theta_0)|^2}{|R_l^u(e^{j\omega},\theta_0)|^2} \Phi^{NS}_{r,opt}(\omega) \qquad \text{(C.1)}$$

Let us note that the ratio in the right hand side of (C.1) can neither be infinite nor zero due to Assumption 4 (in particular, this ratio is equal to one at the frequencies ω where $K_l(e^{j\omega})$ is infinite). Consequently, the spectrum $\Phi_r^S(\omega)$ is well defined at all frequencies ω and is equal to zero at the same frequencies ω as $\Phi^{NS}_{r,opt}(\omega)$ is equal to zero.

If an excitation signal r having the spectrum Φ_r^S (see (C.1)) is used during an experiment where the stealth configuration is implemented, it is clear that the obtained covariance matrix $P_{\theta_l,S}(\Phi_r^S,\theta_0)$ will be equal to the one obtained in the non-stealth case with $\Phi^{NS}_{r,opt}$ (i.e. $P_{\theta_l,NS}(\Phi^{NS}_{r,opt},\theta_0)$). Consequently, an experiment with this spectrum Φ_r^S will satisfy the accuracy constraint (22). Due to (C.1), (A.1) and (A.2), we have also that $J_l^S(\Phi_r^S,\theta_0) = J_l^{NS}(\Phi^{NS}_{r,opt},\theta_0)$ for the individual cost at Node l. Moreover, for $i \neq l$, we have that:

$$|R_i^u(e^{j\omega},\theta_0)|^2 \Phi_r^S(\omega) = \dots$$
$$\dots |\mathcal{N}_i(e^{j\omega},\theta_0)|^2 |T_{0,l}(e^{j\omega}) - T_{init,l}(e^{j\omega})|^2 \Phi_r^S(\omega)$$

$$|R_i^{u,NS}(e^{j\omega},\theta_0)|^2 \Phi^{NS}_{r,opt}(\omega) = \dots$$
$$\dots |\mathcal{N}_i(e^{j\omega},\theta_0)|^2 |T_{0,l}(e^{j\omega})|^2 \tfrac{|R_l^u(e^{j\omega},\theta_0)|^2}{|R_l^{u,NS}(e^{j\omega},\theta_0)|^2} \Phi_r^S(\omega)$$

Consequently, due to Assumption 1, for $i \in \mathcal{P}_l$, we have that $J_i^S(\Phi_r^S,\theta_0) < J_i^{NS}(\Phi^{NS}_{r,opt},\theta_0)$ if $T_{init,l}$ satisfies (23) for all frequencies ω where $\Phi_r^S(\omega) \neq 0$ (and thus for all frequencies ω where $\Phi^{NS}_{r,opt}(\omega) \neq 0$). Moreover, for $i \neq l$ and $i \notin \mathcal{P}_l$, we have that $J_i^S(\Phi_r^S,\theta_0) = J_i^{NS}(\Phi^{NS}_{r,opt},\theta_0) = 0$. Since J is the sum of the individual costs J_i (see Section 3.2), we have thus shown that, in the stealth case, we can find a spectrum Φ_r^S leading to the same accuracy as with $\Phi^{NS}_{r,opt}$, but with a strictly smaller cost. The result of the proposition is therefore proven since, by definition, $J^S(\Phi_r^S,\theta_0) \geq J^S(\Phi^S_{r,opt},\theta_0)$